

# Shahid Ul Islam

Kashmir, India | +917780858125 | <mailto:shahid9664@gmail.com> | [LinkedIn](#) | [Portfolio](#)

## SUMMARY

---

I'm a Data Scientist and Analyst with strong experience in Python, SQL, and machine learning. I've worked on building predictive models, regression analyses, and web-based applications to deliver actionable insights. I'm skilled in data preprocessing, feature engineering, and model evaluation, always focused on turning data into solutions. I'm excited about using AI to tackle real-world problems and help push forward innovative technologies.

## PROJECT EXPERIENCE

---

### **Islamic University of Science & Technology**

#### Breast Cancer Prediction Using Machine Learning

*Awantipora, Pulwama, Jammu and Kashmir*

*Dec 2024 – Jan 2025*

- Developed and implemented a web app using Streamlit that utilizes a predictive model to classify tumors as malignant or benign, aiding in the early detection and diagnosis of breast cancer.
- Conducted extensive preprocessing, including handling missing values, feature scaling, and feature selection, to improve model performance.
- Trained and evaluated classification algorithm, Random Forest Classifier for the best performance.
- Focused on explainability by generating plots to help users understand feature importance in predictions.

**Technologies Used:** Python, NumPy, Pandas, Matplotlib, Streamlit, Plotly, Sklearn, Joblib, Seaborn.

**GitHub:** <https://github.com/Khanz9664/Breast-Cancer-Prediction>

### **Islamic University of Science & Technology**

#### Apple Sales Analysis & Dashboard 2024

*Awantipora, Pulwama, Jammu and Kashmir*

*Jan 2025 – Feb 2025*

- Analyzed and visualized the Apple Sales 2024 dataset to identify key trends and patterns.
- Developed an interactive dashboard for product sales, regional distribution, and revenue insights.
- Applied EDA, statistical techniques, and clustering for actionable insights.
- Built predictive models to forecast sales trends and aid decision-making.
- Enhanced data interpretability using advanced visualization tools.

**Technologies Used:** Python, Pandas, NumPy, Matplotlib, Seaborn, Plotly, Dash, Scikit-learn.

**GitHub:** <https://github.com/Khanz9664/Comprehensive-Data-Analysis-Visualization-of-Apple-Product-Sales>

### **Islamic University of Science & Technology**

#### Medicine Review Categorization

*Awantipora, Pulwama, Jammu and Kashmir*

*Sep 2024 – Dec 2024*

- Designed and developed a web-based application to classify medicine reviews into predefined categories, assisting healthcare professionals and pharmaceutical companies in understanding patient feedback.

- Built the backend logic with a trained classification model using TF-IDF vectorization for text data and label encoding for categorical inputs.
- Integrated the model into a user-friendly Streamlit interface for seamless interaction and predictions.
- Preprocessed data to optimize model accuracy and combined multi-feature inputs into a sparse matrix for efficient processing.
- Focused on scalability, enabling future enhancements like multi-language support and confidence score visualization.

**Technologies Used:** Python, Numpy, Pandas, Streamlit, Scikit-learn, Joblib, TF-IDF, Sparse Matrices

**GitHub:** <https://github.com/Khanz9664/Medicine-Review-Categorization>

## **Indira Gandhi National Open University** Criminal Records Management System

*Srinagar, Jammu and Kashmir*  
*Aug 2022 - Dec 2022*

- Developed a centralized web-based application for managing and analyzing criminal records. The system is designed to streamline operations for law enforcement agencies such as the police, NCRB, and CBI, enabling them to securely manage, update, and retrieve criminal data efficiently.

**Technologies Used:** HTML, CSS, PHP, MySQL.

**GitHub :** <https://github.com/Khanz9664/criminal-records-system>

## **EDUCATION**

---

Islamic University Of Science And Technology.  
Master of Computer Applications (MCA)

*Awantipora, Pulwama, Jammu and Kashmir*  
*August 2023 – Present*

Indira Gandhi National Open University.  
Bachelor Of Computer Applications (BCA)

*Srinagar, Jammu and Kashmir*  
*January 2020 – June 2023*

Govt. Higher Secondary School  
Secondary School Education, Mathematics and Computer Science.

*Kulgam, Jammu and Kashmir*  
*January 2015 – March 2017*

## **LICENSES AND CERTIFICATIONS**

---

### Courses:

- Advanced Learning Algorithms (Coursera)
- Supervised Machine Learning: Regression and Classification (Coursera)
- Professional Diploma in Python Development (Udemy)
- Fundamental Course of Data Architecture 2.0. (Udemy)
- Machine Learning Certification Course for Beginners. (Analytics Vidya)
- Advanced Excel Course. (ELearnMarkets)
- CSS, Bootstrap and Javascript and Python Stack Course. (Udemy)

### Badges:

- Analyzing and Visualizing Data in Looker (Google)
- Introduction To Data Analytics on Google Cloud (Google)
- Introduction To Generative AI (Google)
- Introduction To Large Language Models (Google)
- Introduction To Responsible AI (Google)
- Professional Machine Learning Engineer Study Guide (Google)

## **SKILLS & INTERESTS**

---

### Programming Languages

- Python (Expertise in Pandas, NumPy, Matplotlib, Scikit-learn)
- C, C++ (DSA)
- SQL (Database Design and Query Optimization)
- HTML, CSS, JS

### Data Science & Machine Learning

- Machine Learning Algorithms (Linear Regression, Decision Trees, Random Forest, SVM, ...)
- Data Visualization (Matplotlib, Seaborn)
- Data Analysis and Preprocessing

### Artificial Intelligence & Natural Language Processing

- Knowledge in AI and NLP techniques
- Sentiment Analysis using TextBlob

### Database Management

- Database Design and Management
- Expertise in Relational Database Systems

### Technical Tools

- Jupyter Notebooks
- Git and GitHub for Version Control
- Docker
- VS Code
- Excel
- Google GCP (Big Query & Looker)